



# JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

## Interpretation of Lung cancer using Inception based Google Network

“Rachit Sardar ,Yash Badodekar ,Abhijeet Shimple ,Shivam Raj”

(Students, Department of Computer Engineering, SAE Kondhwa, Pune)

Guided By,

Prof. P. S. Gawali

### I. INTRODUCTION

**Abstract-** Identification of lung disease is the most fascinating exploration zone of analyst's in early days. The proposed framework is intended to identify lung malignancy in untimely stage in two phases. The proposed framework comprises of numerous means, for example, image extraction, pre-processing, binarization, thresholding, Division, highlight extraction, and neural system identification.

In our system we developed Lung Cancer detection system based on machine learning and neural network. It decreases the chances of getting harm to human by early detection of cancer. Presently a few frameworks are proposed and still a large number of them are theoretical plan. Convolutional Neural Network based Classification and location arrangement of lung tumour.

**Keywords-** Convolutional Neural Network, Computed Tomography, Support Vector Machine, and Small Cell Lung Cancer.

Due to large prevalence of smoking and air pollution around the world, lung cancer has become one of the most common and deadly disease in recent decades. It often takes long time to develop and most people are diagnosed with the disease within the age bracket 55 to 65. Early identification and treatment is the best available option for the infected people. Reliable identification and classification of lung cancer requires pathological test, namely, needle biopsy specimen and analysis by experienced pathologists. However, because it involves human judgment of several factors and a combination of experiences, a decision support system is desirable in this case. Recent developments in image processing, pattern recognition, dimensionality reduction and classification methods has paved the way for alternate identification and classification approaches for lung cancer.

In addition to machine learning approaches, deep learning through restricted Boltzmann machine in the form of auto encoders has shown promising success in classification tasks in different domain including acoustics, sentiment classification, and image and text recognition. Motivated by the success of deep learning in relevant fields, a deep learning based classification method is investigated in this work.

## II. LITERATURE SURVEY

### 1. Cancer lung detection on CT scan image using ANN back propagation based gray level co-occurrence matrix feature.

This research focuses on detection of lung cancer utilizing Artificial Neural Network Back-engendering based Gray Level Co-event Matrices (GLCM) highlight. The lung information utilized begins from the Cancer imaging archive Database, information utilized comprised of 50 CT-pictures. CT-picture is gathered into 2 bunches, typical and lung disease. The means of this exploration are: picture pre-processing, locale of intrigue division, highlight extraction, and recognition of lung disease utilizing Neural Network Back-spread. The outcomes demonstrates framework can identify CT-picture of ordinary lung and lung malignancy with exactness of 80%. Early discovery of lung malignant growth will recoup the patient. Instrument used to recognize lung malignant growth is through CT scan (Computed Tomography).

### 2. Lung cancer detection with fusion of CT and MRI image using image processing

The purpose of this paper is to locate the beginning times of lung malignancy and increasingly exact outcome by utilizing distinctive strategies like combination, upgrade and division process. Already the vast majority of the malignant growth location strategies rely upon human experience by watching the picture of CT-filter. It will be a bogus discovery of lung malignant growth arrange. Utilizing Image Processing we can rapidly and precisely distinguish tumour of malignant growth. Utilizing Image Processing viable procedures we gather data from complex restorative pictures. In combination procedure, the critical highlights of various unique pictures are consolidated together to acquire the required data in a Fused Image. In medicinal application there are different plans to enhance the substance of picture shape CT and MRI like CT picture examines the denser tissues and MRI filters the delicate tissues, so by joining pertinent data of the two pictures, we get proper data of melded picture. This procedure additionally enhances the nature of the melded picture.

### 3. Segmentation and analysis of CT chest images for early lung cancer detection

In this paper an improved technique for Hopfield Artificial Neural Network Classifier show is proposed to section extricated lung locales from human chest Computer Tomography pictures. The pictures are procured utilizing Computer Tomography imaging methods from typical subjects and others as possibility for lung malignant growth determination. A blend of bit-planes of every pixel are utilized to upgrade edges' recognition of lung area flaps. Three indicative guidelines are confirmed too characterized channels of applicant malignant locales from the status of possibility to false or genuine positive status.

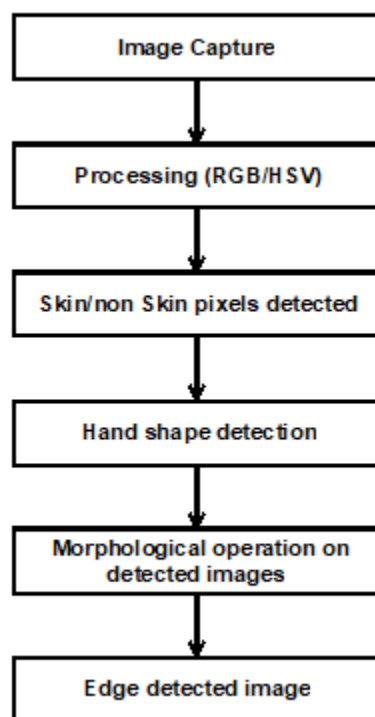
### 4. Lung cancer classification using deep learned features on low population dataset

In this paper, we propose a novel neural-arrange based calculation, which we allude to as entropy debasement technique (EDM), to recognize little cell lung malignant growth (SCLC) from processed tomography (CT) pictures. This exploration could encourage early identification of lung malignant growths. The preparation information and testing information are high-goals lung CT examines given by the National Cancer Institute. We chose 12 lung CT filters from the library, 6 of which are for sound lungs, and the rest of the 6 are examines from patients with SCLC. We arbitrarily take 5 filters from each gathering to prepare our model, and utilized the staying two outputs to test. Our calculations accomplishes a precision of 77.8%.

### 5. Small-cell lung cancer detection using a supervised ML algorithms

In this work, a profound auto encoder order system is proposed which initially adapts profound highlights and after that prepares a fake neural system with these educated highlights. Test results demonstrate the profound educated classifier beats every single other classifier when prepared with all properties and same preparing tests. It is likewise exhibited that the execution enhancement is factually huge.

### III.EXISTING SYSTEM APPROACH



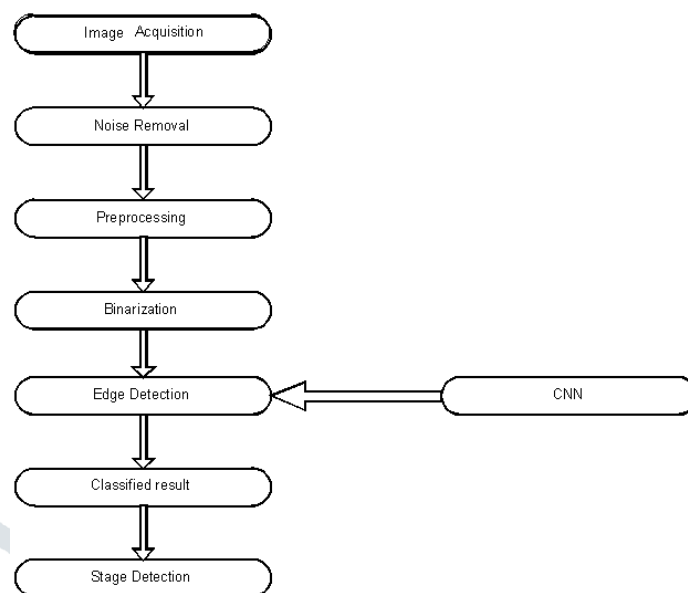
**Fig.1 Block Diagram of Existing System**

The existing system presented uses Support Vector Machine Classifier and Watershed Segmentation of cancer detection. Also uses ANN back-propagation based on Gray Level Co-occurrence Matrix features. Which methods not fully accurate prediction methods thus we have to implement our system to overcome this drawbacks.

#### Disadvantages

1. Due to large prevalence of smoking and air pollution lung cancer has become one of the most deadly disease in recent decades.
2. System needed by improving the preprocessing process, image segmentation, feature extraction, and learning process.

### IV.PROPOSED SYSTEM APPROACH



**Fig.2 Block Diagram of Proposed System**

We propose a novel Lung detection and Stage prediction mechanism is proposed which first learns deep features and then trains an artificial neural network with these learned features. Experimental results show the deep learned classifier outperforms all other classifiers when trained with all attributes and same training samples. It is also demonstrated that the performance improvement is statistically significant. Classification of lung cancer using a low population, high dimensional data-set is challenging due to insufficient samples to learn an accurate mapping among features and class labels. Current literature usually handles this task through handcrafted feature creation and selection. Deep learning is found to be able to identify the underlying structure of data through the use of CNN and other techniques.

#### Advantages

1. This shows that application of machine learning has the potential to significantly detect and classify with almost high accuracy for the low population...
2. High dimensional lung cancer data-set without requiring any hand-crafted, case specific features.
3. High processing speed enhanced CNN classifier model.

## V.SYSTEM ARCHITECTURE

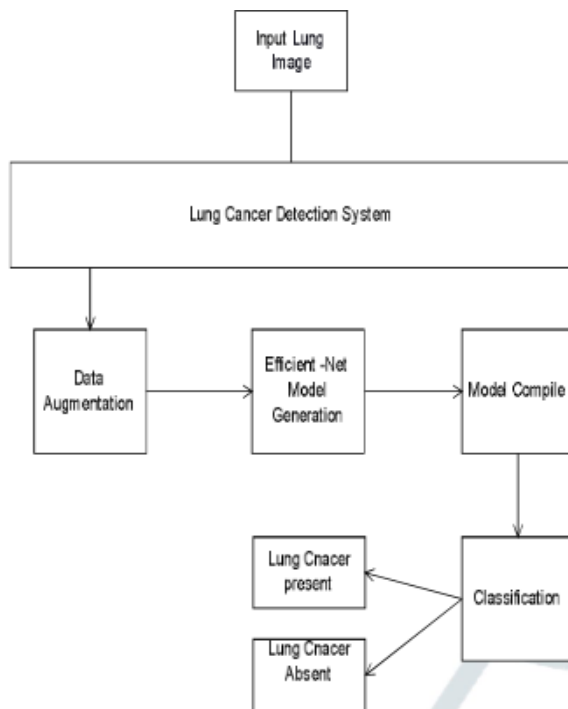


Fig.3 System Architecture

### METHODOLOGY USED IN PROPOSED SYSTEM:

#### A) Image Filtering:-

Filtering is a technique to modify or enhance the image, i.e. to highlight certain features or remove other features. It includes smoothing, sharpening. And edge enhancement. Image filtering algorithms generate an output pixel by observing the neighbourhood of the input pixel in an image. Image filtering algorithms are used to remove different types of noise from the image.

#### B) Segmentation:-

Image segmentation is the way toward apportioning an advanced picture into various portions (sets of pixels). All pixels in an area share a typical property. The objective is to disentangle and change the portrayal of the picture into something that is increasingly important and less demanding to break down.

#### C) Edge Detection:-

Edge defines the boundaries between regions in an image which helps in object detection. There are

many edge detection operators and algorithms available. Edge Detection Operators and Algorithms used in our research like Canny Edge Detection method.

#### D) Neural Network:-

Brain-inspired systems are to replicate how humans learn. Consist of input, hidden and output layers that transform the input into something that the output layer can use. Excellent for finding patterns which are complex for humans to extract and teach the machine to recognize. ANN gathers their knowledge by detecting the patterns and relationships in data and learn (or are trained) through experience, not from programming.

## VI. MATHEMATICAL MODEL

- **System Description:**  $S = \{I, F, O\}$  S be the lung cancer detection and stage prediction system.

#### INPUT:

- $F = F_1, F_2, F_3 \dots F_N$  Functions to be execute modules
- $I = C_1, C_2, C_3 \dots$  input of systems lung cancer images
- $O = R_1, R_2 \dots R_n$  prediction results in the form of stages
- I=result access by User

#### F:

- **F1**=Image processing applied on Lung Cancer images
- **F2**=feature extraction from images
- **F3**=Stage detection results

#### O:

- $R_1$  = model creation from training.
- $R_2$  = model based image testing

#### Success:

1. High accuracy achieved by using all type of MRI image dataset.
2. User gets result very fast according to their needs.

**Failures:**

1. Huge database can lead to more time consumption to get the information.
2. Hardware failure.
3. Software failure.

**Mathematical Model in Equation format****Notation**

Where,

- M= Set of all entities.
- LCIT1= Lung cancer images type 1
- LCIT2= Lung cancer images type 2
- LCITN= Lung cancer images type N
- TLCI=Total Lung cancer images

For calculate total number of users by following equation 1

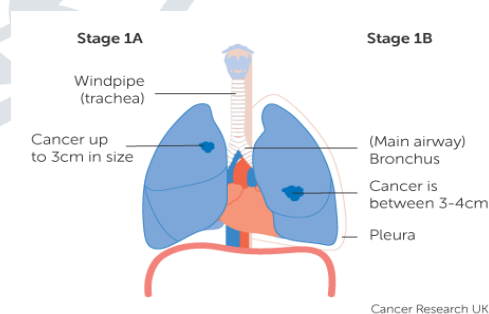
**Total number lung cancer images = Total number lung cancer images type 1+ = Total number lung cancer images type 2+.....+ = Total number lung cancer images type N**

$$TLCI = \sum LCIT1 + \sum LCIT2 + \dots + \sum LCITN \dots \text{equation 1}$$

**VII. COMPARATIVE RESULT****Stages of Lung Cancer:-**

- **Stage I:** Stage 1 is a part of number staging system and means your cancer is small. It hasn't spread to your lymph nodes and other distinct organ.
  - they are of two types :
- Type A:- Means that the cancer is 2 cm or smaller.
- Type B :-Means that the cancer is between 2 to 3 cm

| Sr No. | Types                                        | Number of images |
|--------|----------------------------------------------|------------------|
| 1      | Lung cancer testing dataset images<br>type 1 | 210              |
| 2      | Lung cancer testing dataset images<br>type 2 | 175              |
| 3      | Lung cancer testing dataset images<br>type 3 | 385              |
| 4      | Lung cancer images testing dataset<br>type 4 | 230              |

**Table 8.1: Model Testing****Fig. Lung Cancer Stage 1**

- ❑ **STAGE II:** These tumours are larger than the stage 1 cancers and start to spread to nearby lymph nodes or other structures.
- ❑ About 30% of the lungs cancer are diagnosed at these stage, and the treatment success rate is higher than the later stage.



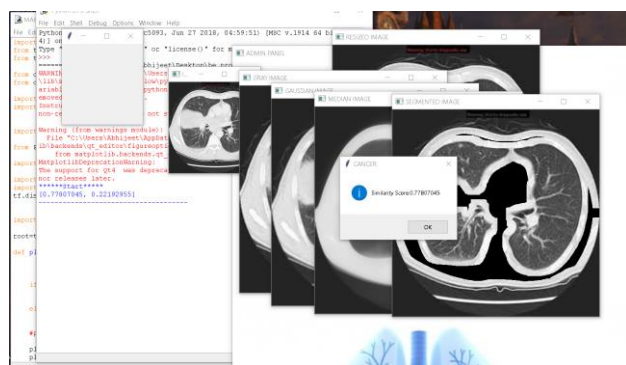


Fig. Lung cancer Stage 2

- Stage III : In this stage the cancer is more than 3 cm wide and is spread in the lymph nodes and other parts outside the lungs. in this stage the affected lymph nodes are restricted to the same side of the body as the cancer.
- Stage IV : The cancer is metastasized or spread beyond the lungs into other areas of the body.

| Primary Cancer(T) | Criteria                                           |
|-------------------|----------------------------------------------------|
| T1                | < 3cm in diam; T 1a <= 2cm; T 1b > 2cm <= 3cm      |
| T2                | > 3cm <= 5cm; T 2a > 3cm <= 4cm; T 2b > 4cm <= 5cm |
| T3                | >5cm<=7cm                                          |
| T4                | Any Size greater than above                        |

Table. Lung cancer stage evaluation

From above data, as shown in graph 9.1, the total numbers of images of lung cancer type 1 were 210, total numbers of images of lung cancer type 2 were 175, total numbers of images of lung cancer type 3 were 385 and total numbers of images of lung cancer type 3 were 230.

### VIII. CONCLUSION

This shows that application of machine learning has the potential to significantly detect and classify with almost accuracy for the low population, high dimensional lung cancer dataset without requiring any hand-crafted, case specific features.

### VII. FUTURE WORK

For future work, we can invent this technique on large dataset images. Increasing the number of images will be used for the process; can improve the accuracy of our system.

### VIII. ACKNOWLEDGMENT

This work is supported in Lung Cancer Detection System for medical research and development of any state in India. Authors are thankful to Faculty of Engineering and Technology (FET), Savitribai Phule Pune University, Pune for providing the facility to carry out the research work.

### IX. REFERENCES

- [1] Lilik Anifah, Haryanto, Rina Harimurti, "Cancer lung detection on CT scan image using ANN backpropagation based gray level co occurrence matrix feature." 978-1-5386-3172-0/17/ 2017 IEEE
- [2] Prof. Anuradha Deshpande, Dhanesh Lokhande, "Lung cancer detection with fusion of CT and MRI image using image processing." (IJARCET) Volume 4 Issue 3, March 2015
- [3] Rachid Sammouda, "Segmentation and analysis of CT chest images for early lung cancer detection." Global Summit on Computer & Information Technology 978-1-5090-2659-3/17 2017 IEEE
- [4] Qing Wu, Wenbing Zhao, "Small-cell lung cancer detection using a supervised ML algorithms." International Symposium on Computer Science and Intelligent Controls 978-1-5386-2941-3/17 2017 IEEE
- [5] Abbas Khosravi, Amin Khatami, "Lung cancer classification using deep learned features on low population dataset." Canadian Conference on Electrical and Computer Engineering (CCECE) 978-1-5090-5538-8/17 2017 IEEE

[6] Sanyogita Bobade, Pratik Deshpande, Abhinav Deshpande, Aditi Deshmukh, Shalini Wankhade, “Detecting Lung Cancer from Live Dataset Using CNN.” (IJIRSET) Volume 9, Issue 9, September 2020

[7] Shalini Wankhade, Dr. Vigneshwari S, “Lung Nodule Detection and Classification based on Feature Merging and Genetic Algorithm.” International Journal of Future Generation Communication and Networking Volume 13, Issue 4, 2020

### Authors

**First Author** - Arya Honraopatil, B.E. Computer, Sinhgad Academy of Engineering, [aryahonraopatil@gmail.com](mailto:aryahonraopatil@gmail.com)

**Second Author** - Rutuja Bhalerao, B.E. Computer, Sinhgad Academy of Engineering, [bhaleraorutuja2000@gmail.com](mailto:bhaleraorutuja2000@gmail.com)

**Third Author** - Prachi Chavan, B.E. Computer, Sinhgad Academy of Engineering, [prachichavan483@gmail.com](mailto:prachichavan483@gmail.com)

**Fourth Author** - Nikita Desai, B.E. Computer, Sinhgad Academy of Engineering, [nikitadesai0305@gmail.com](mailto:nikitadesai0305@gmail.com)

**Fifth Author** – Prof. Shalini Wankhade, M.E. Computer, Sinhgad Academy of Engineering, [svwankhade.sae@sinhgad.edu](mailto:svwankhade.sae@sinhgad.edu)

**Correspondence Author** – Arya Honraopatil, [aryahonraopatil@gmail.com](mailto:aryahonraopatil@gmail.com), 8380077039

